



Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at <http://about.jstor.org/participate-jstor/individuals/early-journal-content>.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact support@jstor.org.

ANALYSIS OF THE STANDARD OF LIVING IN THE
DISTRICT OF COLUMBIA IN 1916.*

BY WILLIAM F. OGBURN.

At the close of the year 1916, the U. S. Bureau of Labor Statistics collected over two thousand budgets of families living in the District of Columbia showing details of expenditures for the calendar year 1916. An analysis of these budgets was published in the "Monthly Review of the U. S. Bureau of Labor Statistics" for October and November, 1917. It is the purpose of this paper to show a somewhat more detailed analysis of this study in order to bring out more nearly complete information concerning the lives and habits of these families, than appears in a brief analysis. In doing this, it is also the purpose of this paper, to use a method which not only presents rather full information with considerable brevity, but generalizes from the sample in such a form as to permit a large number of comparisons, and to make a number of detailed estimates and predictions.

The budgets were collected in such a manner as to be representative of the community as a whole for families of low and moderate incomes. They were from the black and white races and included a variety of nationalities and occupations. There was no limitation as to size and constituency of the family. Families with incomes over \$2,000 a year for others than wage earners were excluded; as were also families in which the husband earned less than 75 per cent. of the total income. The budgets were collected in a house-to-house canvas by agents. Of these 2,110 budgets 200 were selected for the purpose of this analysis. The selection was of white families comprising at least a man and wife living in rented houses without lodgers, boarders or adults, other than children. Also only those schedules were chosen which contained full detailed data on expenditures for sundries. This selection yielded only 200 cases. About 90 different occupations were represented, distributed among the unskilled, the semi-skilled, and the skilled trades, as well as among other occupations as clerks and salesmen.

* Paper read at the Eightieth Annual Meeting of the American Statistical Association.

What we really want to know about such budgetary studies is how the expenditure for various items varies with the size of the income and with the size of the family. The problem of analysis consists, then, of classifying the various items of expenditure by income and by size of the family. The items of expenditure studied in this analysis are food, rent, fuel and light, clothing, sundries, husband's clothing, wife's clothing, children's clothing, insurance, furniture and furnishings, contribution for religious purposes, education, amusement, liquor and tobacco and sickness. Various relationships may also be shown between these items of expenditure. A list of the variables, their definition and designation, used in these relationships follows.

DEFINITION OF VARIABLES.

X_1 = Annual income of family for a year in dollars, and includes earnings from garden, earnings of wife, if any, and all other income, including gifts.

X_2 = Deficit or surplus, and is the difference between income and expenditure. Expenditure includes the total for the same year for which income is computed, and includes unpaid bills, such as doctors' bills. A deficit is listed with a minus sign and a surplus with a plus sign.

X_3 = Size of family, as measured in units of an adult male, based on food consumption. The U. S. Bureau of Labor Statistics scale is used, which rates a woman as 0.9 of an adult male, child (11 to 14) as 0.9, child (7 to 10) as 0.75, child (4 to 6) as 0.4, and child (to 3) as 0.15. A further discussion of this measure of size of family is found in a later paragraph.

X_4 = Percentage of total expenditure spent for *food* in a year for a whole family, and includes confectioneries, lunches taken out, and ice.

X_5 = Percentage of total expenditure spent for *rent* in a year for the whole family. If two or more houses were lived in, the rents for each fraction of a year were totaled. Owned houses and houses in which lived lodgers or boarders with the family were not included. Rent does not include fuel and light (except in four cases).

X_6 = Percentage of total expenditure spent for *fuel and light* in a year for the whole family.

X_7 = Percentage of total expenditure spent for *clothing* for the whole family. No family includes others than man, wife and children, and at least every family had a man and wife.

X_8 = Percentage of total expenditure spent for *sundries*. Sundries means all expenditures other than expenditures for food, rent, fuel and light, and clothing. It includes insurance, dues to organizations, payments for religion, amusement, education, sickness, furniture and furnishings, and expenses for various incidentals and exigencies.

X_9 = Percentage of total expenditure spent on *husband's clothing*.

X_{10} = Percentage of total expenditure spent on *wife's clothing*.

X_{11} = Percentage of total expenditure spent on *children's clothing*.

X_{12} = Percentage of total expenditure spent for *insurance* of all kinds, including fraternal, industrial and "old line" insurance.

X_{13} = Percentage of total expenditure that went to institutions or organizations of *religion*.

X_{14} = Percentage of total expenditure spent for *furniture and furnishings*. In cases of furniture bought on the installment plan, the installments which are paid and are due during the year are included.

X_{15} = Percentage of total expenditure spent on *education*. Educational expenditures include in this study essentially, newspapers, books, magazines and school expenses of children of high school age. Education is a somewhat inaccurate descriptive term for the items of expenditure included here.

X_{16} = Percentage of total expenditure spent for *amusement*, including movies, theatres, dances, pool, etc. Amusement generally is a much broader term than is here suggested by these particular items of expenditure.

X_{17} = Percentage of total expenditure for *liquor and tobacco*. At this time it was legal to sell liquor in Washington.

X_{18} = Percentage of total expenditure spent for *sickness*. This includes unpaid doctors' bills incurred during the year, also medicines, drugs, nurse and hospital expenses. The meaning of the correlations involving amount spent for sickness is not clear for the amount spent for sickness does not necessarily indicate the number of days of sickness.

X_{19} = Total clothing expenditure in multiples of the amount spent for the husband's clothing, and is found by dividing the expenditure for man's clothing into the total expenditure of clothing. These ratios are studied for the purpose of supplementing the information gathered from the indexes of equivalent adult males in food consumption as to validity of these figures as indexes of the size of the family.

X_{20} = *Cost per adult male per day* in food consumption, and is determined by dividing the total yearly family food bill by 365 and then by dividing by the number of equivalent adult males in food consumption in the family.

Table I shows how the data may be classified. In this table, the percentage of expenditure for food is classified according to families of various incomes. This table shows therefore the relationship between income and percentage of expenditure for food. Similarly the relationship between income and the various other items of expenditure have been analyzed, the tables not being presented here. Also the percentages of expenditure for various items of the family budget have been classified, according to size of family in similar tables, and other tables have been constructed showing various relationships.

TABLE I.

RELATIONSHIP BETWEEN SIZE OF ANNUAL INCOME AND PERCENTAGE OF TOTAL ANNUAL EXPENDITURE SPENT FOR FOOD.

	500-	600-	700-	800-	900-	1000-	1100-	1200-	1300-	1400-	1500-	1600-	1700-	1800-	1900-	2000-	2100-	2200-	
20-	..	..	..	..	..	..	..	..	..	1	..	..	..	..	..	..	..	..	1
22-	..	..	1	..	..	..	..	..	..	1	1	..	..	..	..	..	..	..	4
24-	..	..	..	..	..	..	..	..	..	..	..	1	1	..	..	1	..	..	4
26-	..	..	..	1	..	..	..	1	..	..	2	..	..	..	1	..	..	..	5
28-	..	..	..	..	1	1	1	1	..	2	..	..	..	..	..	..	..	1	8
30-	..	1	..	..	..	1	..	..	..	2	2	1	..	..	..	..	..	..	7
32-	..	..	..	1	..	1	2	2	..	1	1	1	1	1	..	..	..	..	17
34-	..	..	2	2	..	1	..	1	4	3	3	3	1	..	..	..	..	..	13
36-	..	1	..	2	1	3	3	4	1	2	3	..	..	..	..	..	..	..	13
38-	..	1	1	2	4	3	2	2	1	3	3	..	..	..	..	..	..	..	22
40-	..	1	1	5	1	2	5	4	3	3	1	1	1	..	..	..	..	..	26
42-	1	1	1	3	3	..	1	3	1	1	1	..	..	..	..	..	..	..	16
44-	..	1	3	1	2	1	..	1	..	2	1	2	..	..	..	..	..	..	14
46-	2	1	5	3	1	..	1	..	3	1	..	..	1	..	..	..	..	..	18
48-	1	1	4	1	2	..	..	1	..	..	..	..	..	..	..	..	..	..	10
50-	..	2	3	2	2	1	..	1	1	..	..	..	..	1	..	..	..	..	13
52-	..	1	1	..	..	..	1	..	..	..	..	..	..	..	..	..	..	..	3
54-	..	..	2	2	..	..	..	..	..	..	..	..	..	..	..	..	..	..	4
56-	..	1	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	2
	4	12	24	25	18	13	13	21	13	22	17	8	4	2	2	1	..	1	200

Only one of these tables showing relationships is included in this report, but in Table II an abbreviated or condensed description of these various tables of relationships is shown. Table II contains a list of the correlation coefficients measuring these relationships. Thus we see that the correlation coefficient between income and percentage of expenditure for food is high, $r_{1.4} = -0.513$, while the correlation coefficient between income and percentage of expenditure for rent is low, $r_{1.5} = -0.10$. These coefficients tell us that the size of the income affects only slightly percentages of expenditure for rent, while increasing the income markedly diminishes the percentage of expenditure for food. And so the various relationships can be interpreted from the correlation coefficients.

TABLE II.
CORRELATION COEFFICIENTS.

$r_{2.1} = +0.354$	$r_{2.8} = -0.196$
$r_{3.1} = +0.047$	$r_{4.8} = +0.256$
$r_{4.1} = -0.513$	$r_{5.8} = -0.075$
$r_{5.1} = -0.10$	$r_{6.8} = +0.028$
$r_{6.1} = -0.227$	$r_{7.8} = +0.165$
$r_{7.1} = +0.351$	$r_{8.8} = -0.255$
$r_{8.1} = +0.419$	$r_{9.8} = -0.286$
$r_{9.1} = +0.229$	$r_{10.8} = +0.438$
$r_{10.1} = +0.368$	$r_{11.8} = +0.692$
$r_{11.1} = +0.071$	$r_{12.8} = +0.011$
$r_{12.1} = -0.022$	$r_{13.8} = -0.0284$
$r_{13.1} = +0.175$	$r_{14.8} = -0.115$
$r_{14.1} = +0.10$	$r_{15.8} = +0.0038$
$r_{15.1} = +0.304$	$r_{16.8} = -0.147$
$r_{16.1} = +0.265$	$r_{17.8} = -0.029$
$r_{17.1} = -0.122$	$r_{18.8} = -0.078$
$r_{18.1} = -0.036$	$r_{19.8} = +0.405$
$r_{19.1} = -0.046$	$r_{20.8} = -0.639$
$r_{20.1} = +0.391$	

Although the correlation coefficients show the results of the classification of items of expenditure according to income and according to size of family, perhaps a better idea can be had by expressing these relationships in equations rather than in coefficients. These equations can best be explained by observing the plotting in Diagram I.* In this chart we see from the location of the circular dots the average amount of deficit or surplus, according to size of income. The procession of circular dots shows that for very low incomes there is a large deficit and that as the income increases the deficit becomes less

* This diagram was not available for the QUARTERLY, but will be published in another edition of the Author's work.—ED.

and less, and soon there is a surplus which increases with the income. There is considerable fluctuation of these circular dots or means, due to the fact that the sample includes only 200 cases. If the sample had been very large the procession of the means would have approximated with less fluctuation a straight line. We can determine the equation to the straight line which the dots follow. It is $X_2 = -166.45 + 0.144 X_1$ and the plotting of it is shown in Diagram I. This "regression" equation expresses the relationship between size of income and the amount of deficit or surplus in somewhat more interpretable form than does the correlation coefficient. Thus the equation tells us that for every increase of \$100 in income a year the amount of deficit is reduced \$14 or the surplus is increased \$14. We can also substitute in this equation any income (between \$800 and \$1,800 a year) and know what on the average is the amount of deficit or surplus of a family living in the District of Columbia in 1916. These results can also be read from the plotting of the line in the diagram. Thus when the income was \$1,500 there was on the average a surplus of \$50 and when the income was \$1,155 there was, on the average, neither deficit nor surplus. Similarly the other equations in Table III show the relationship existing between the other variables.

TABLE III.
REGRESSION EQUATIONS.

$X_2 = -166.45 + 0.144X_1$	$X_2 = 81.09 - 24.22X_3$
$X_3 = -3.23 + 0.00016X_1$	$X_4 = 34.31 + 1.71X_3$
$X_4 = 53.08 - 0.0113X_1$	$X_5 = 20.66 - 0.297X_3$
$X_5 = 21.14 - 0.0013X_1$	$X_6 = 5.52 + 0.05X_3$
$X_6 = 7.19 - 0.0013X_1$	$X_7 = 9.03 + 0.496X_3$
$X_7 = 6.70 + 0.0035X_1$	$X_8 = 30.06 - 1.80X_3$
$X_8 = 12.70 + 0.0098X_1$	$X_9 = 5.30 - 0.47X_3$
$X_9 = 2.21 + 0.0013X_1$	$X_{10} = 4.73 - 0.635X_3$
$X_{10} = 0.50 + 0.0018X_1$	$X_{11} = -0.80 + 1.57X_3$
$X_{11} = 3.94 + 0.00053X_1$	$X_{12} = 4.03 + 0.024X_3$
$X_{12} = 4.29 - 0.00016X_1$	$X_{13} = 1.01 - 0.026X_3$
$X_{13} = 0.30 + 0.00054X_1$	$X_{14} = 4.07 - 0.34X_3$
$X_{14} = 1.81 + 0.00096X_1$	$X_{15} = 0.77 + 0.0023X_3$
$X_{15} = 0.10 + 0.00059X_1$	$X_{16} = 1.44 - 0.129X_3$
$X_{16} = 0.12 + 0.00077X_1$	$X_{17} = 1.97 - 0.055X_3$
$X_{17} = 2.65 - 0.00076X_1$	$X_{18} = 3.85 - 0.24X_3$
$X_{18} = 3.44 - 0.00036X_1$	$X_{19} = 0.81 + 0.85X_3$
$X_{19} = 4.08 - 0.00032X_1$	$X_{20} = 0.62 - 0.069X_3$
$X_{20} = 0.23 + 0.00014X_1$	

Observation of the list of equations in Table III shows them to be equations of straight lines. In other words, we have

assumed that the circular dots or the means, if the sample had been sufficiently large, would have fallen on a straight line. Of course, there is no *a priori* reason why the means should follow a straight line rather than a curved line, or any other kind of line. Observation, however, of the plotting of the means of the various tables of relationships seems to show that they follow approximately a straight line, at least for incomes from \$800 to \$1,800. The number of cases of incomes beyond these limits are quite small and the scattering for the very low and very high incomes is quite marked. The present study is therefore most accurate for families with incomes between \$800 and \$1,800. In some tables the means for very high incomes seem not to follow a straight line. Thus in Diagram I in the case of size of deficit or surplus, according to size of income, perhaps the line best fitting the means would be a straight line up to \$1,700 and then curving upward slightly. In most of the relationships studied, though, the straight line is a better fit than is true in the case of the data in Diagram I. However, in this diagram the straight line gives a considerably better description than the rough plotting of the means. Thus according to the means, when the income is \$1,150 there is a deficit of \$61, but according to the equation of the line there is no deficit or surplus when the income is \$1,150. In other words, the true amount of deficit or surplus is more nearly indicated by the line than by the means, because the means fluctuate a good deal for the reason that the sample is small. So that even if the straight line is not a perfect fit, the fit is sufficiently good to give a good description. This statement is shown to be true by the following list of tests of "goodness of fit," shown in Table IV. A customary measure of the extent of divergence of the means from a straight line is the difference between the correlation ratio and the correlation coefficient, or, better still, the difference between their squares, which is usually written as ζ . If the ζ is large the fit is not very good. The size of the ζ can be estimated by comparison with its error. If three times the error is less than ζ , it is fairly certain that the straight line is not a good fit. So by comparing in Table IV the ζ 's with their errors, an estimate may be made of the "goodness of fit" of a straight line to the means, for each of the relationships studied. The errors of ζ were derived from the formula, $\epsilon_{\zeta} = 2\sqrt{\frac{\zeta}{n}}$.

TABLE IV.

DIFFERENCES BETWEEN THE CORRELATION RATIO SQUARED AND THE CORRELATION COEFFICIENT SQUARED AND THE ERROR OF THIS DIFFERENCE.

$\zeta_{2-1} = 0.1905$	$\epsilon_{2-1} = 0.0616$	$\zeta_{2-2} = 0.090$	$\epsilon_{2-2} = 0.042$
$\zeta_{3-1} = 0.0633$	$\epsilon_{3-1} = 0.0356$	$\zeta_{4-2} = 0.053$	$\epsilon_{4-2} = 0.033$
$\zeta_{4-1} = 0.0338$	$\epsilon_{4-1} = 0.026$	$\zeta_{5-2} = 0.1026$	$\epsilon_{5-2} = 0.045$
$\zeta_{5-1} = 0.037$	$\epsilon_{5-1} = 0.0272$	$\zeta_{6-2} = 0.059$	$\epsilon_{6-2} = 0.034$
$\zeta_{6-1} = 0.043$	$\epsilon_{6-1} = 0.0294$	$\zeta_{7-2} = 0.098$	$\epsilon_{7-2} = 0.044$
$\zeta_{7-1} = 0.0574$	$\epsilon_{7-1} = 0.0338$	$\zeta_{8-2} = 0.059$	$\epsilon_{8-2} = 0.034$
$\zeta_{8-1} = 0.0596$	$\epsilon_{8-1} = 0.034$	$\zeta_{9-2} = 0.127$	$\epsilon_{9-2} = 0.052$
$\zeta_{9-1} = 0.1554$	$\epsilon_{9-1} = 0.056$	$\zeta_{10-2} = 0.065$	$\epsilon_{10-2} = 0.036$
$\zeta_{10-1} = 0.0892$	$\epsilon_{10-1} = 0.042$	$\zeta_{11-2} = 0.0409$	$\epsilon_{11-2} = 0.028$
$\zeta_{11-1} = 0.0717$	$\epsilon_{11-1} = 0.038$	$\zeta_{12-2} = 0.0834$	$\epsilon_{12-2} = 0.04$
$\zeta_{12-1} = 0.0718$	$\epsilon_{12-1} = 0.038$	$\zeta_{13-2} = 0.0699$	$\epsilon_{13-2} = 0.036$
$\zeta_{13-1} = 0.0605$	$\epsilon_{13-1} = 0.034$	$\zeta_{14-2} = 0.0847$	$\epsilon_{14-2} = 0.04$
$\zeta_{14-1} = 0.0639$	$\epsilon_{14-1} = 0.036$	$\zeta_{15-2} = 0.0324$	$\epsilon_{15-2} = 0.0254$
$\zeta_{15-1} = 0.0474$	$\epsilon_{15-1} = 0.0308$	$\zeta_{16-2} = 0.029$	$\epsilon_{16-2} = 0.024$
$\zeta_{16-1} = 0.1349$	$\epsilon_{16-1} = 0.052$	$\zeta_{17-2} = 0.0407$	$\epsilon_{17-2} = 0.028$
$\zeta_{17-1} = 0.022$	$\epsilon_{17-1} = 0.021$	$\zeta_{18-2} = 0.0315$	$\epsilon_{18-2} = 0.025$
$\zeta_{18-1} = 0.0816$	$\epsilon_{18-1} = 0.04$	$\zeta_{19-2} = 0.042$	$\epsilon_{19-2} = 0.029$
$\zeta_{19-1} = 0.0288$	$\epsilon_{19-1} = 0.024$	$\zeta_{20-2} = 0.0775$	$\epsilon_{20-2} = 0.0394$
$\zeta_{20-1} = 0.079$	$\epsilon_{20-1} = 0.04$		

The correlation coefficients and the regression equations may at times give misleading information, for the reason that they measure the relationship between only two variables, whereas there may be a very important third variable related, which must be taken into consideration before we can get the real meaning. For instance, we see from the fact that $r_{3-2} = -0.196$, in Table II, that the larger the family the less the surplus or the greater the deficit. But we know that the size of the income has a good deal to do with whether there is a surplus or a deficit. What we really want to know is what is the degree of correlation between size of family and deficit or surplus for a constant income. The coefficient r_{3-2} doesn't tell us anything about the income; but we can find out the relationship between size of family and deficit or surplus with a constant income, by using the method of partial correlation. Using this method we find ${}_1r_{3-2} = -0.228$, in Table V, which means that with a constant income the correlation between income and deficit or surplus is somewhat greater than when we measure the correlation, irrespective of whether the income is constant.

Taking another illustration, we see from the fact that $r_{3-15} = +0.0038$, that the larger the family the more is spent for such educational purposes as newspapers, magazines and

books. But we also know that the larger the income the more is spent for such purposes. Now if larger families have larger incomes, the correlation $r_{3.15}$ instead of showing that larger families spend more for educational purposes *may* mean really that families with larger incomes spend more for educational purposes, because large families have larger incomes. This seems to be the case for $r_{1.3}$ (income and size of family) equals $+0.047$, and $r_{1.15}$ (income and education) equals $+0.0304$, and the partial correlation coefficient $r_{3.15}$, a correlation between size of family and percentages of expenditures for educational purposes with a constant income, equals -0.011 . In other words, with the income unchanged, an increase in family is accompanied with a smaller percentage of expenditures for newspapers, magazines and books, which is the opposite of what seemed to be the case when we considered the relationship between income and expenditure for these purposes, considering income as a factor ($r_{3.15} = +0.0038$). The correlations just considered are very small but they illustrate the idea in partial correlation.

From these illustrations it is seen that we really want to know the relationships between income and expenditure for various purposes, with the size of the family kept constant, and also the relationships between size of family and expenditure for various purposes, with the income kept constant. The partial coefficients of correlation in Table V give us this information, and are therefore less misleading in information than the simple correlation coefficients.

The part of the list of partial coefficients of correlation between income and various items of expenditure with the size of the family kept constant, is a sort of chart showing the satiability of human wants. Where the sign of the correlation coefficient is negative the desires seem to be more quickly satisfied, while the large positive correlations indicate the desires that are capable of somewhat greater expansion. Thus the desire for food is not capable of very great expansion and $r_{1.4} = -0.513$. This is not so true of the desire for clothes and in this case $r_{1.7} = +0.351$.

Also the latter half of the list of partial coefficients of correlations, showing correlations between size of family and various expenditures with an income kept constant, is a very

TABLE V.
PARTIAL CORRELATION COEFFICIENTS.

$r^{f\ 2-1} = +0.371$	$r^{f\ 8-3} = -0.302$
$r^{f\ 4-1} = -0.544$	$r^{f\ 9-3} = -0.305$
$r^{f\ 5-1} = -0.0965$	$r^{f\ 10-3} = -0.49$
$r^{f\ 6-1} = -0.228$	$r^{f\ 11-3} = +0.6925$
$r^{f\ 7-1} = +0.349$	$r^{f\ 12-3} = -0.012$
$r^{f\ 8-1} = +0.447$	$r^{f\ 13-3} = -0.0375$
$r^{f\ 9-1} = +0.253$	$r^{f\ 14-3} = -0.121$
$r^{f\ 10-1} = +0.432$	$r^{f\ 15-3} = -0.011$
$r^{f\ 11-1} = +0.0542$	$r^{f\ 16-3} = -0.165$
$r^{f\ 12-1} = -0.022$	$r^{f\ 17-3} = -0.242$
$r^{f\ 13-1} = +0.177$	$r^{f\ 18-3} = -0.076$
$r^{f\ 14-1} = +0.106$	$r^{f\ 19-3} = +0.408$
$r^{f\ 15-1} = +0.305$	$r^{f\ 20-3} = -0.7157$
$r^{f\ 16-1} = +0.275$	$1-8^{f\ 18-5} = -0.182$
$r^{f\ 17-1} = -0.121$	$1-8^{f\ 12-16} = +0.239$
$r^{f\ 18-1} = -0.033$	$1-8^{f\ 12-18} = -0.073$
$r^{f\ 19-1} = -0.03$	$1-8^{f\ 13-16} = -0.081$
$r^{f\ 20-1} = +0.549$	$1-8^{f\ 13-15} = -0.0822$
$r^{f\ 2-3} = -0.228$	$1-8^{f\ 15-16} = +0.054$
$r^{f\ 4-3} = +0.327$	$1-8^{f\ 16-18} = -0.121$
$r^{f\ 5-3} = -0.0708$	$1-8^{f\ 17-18} = -0.138$
$r^{f\ 6-3} = +0.039$	$1-8^{f\ 18-20} = +0.046$
$r^{f\ 7-3} = +0.159$	

good indication of what happens when the family increases and the income does not. The negative coefficients of correlations show the particular curtailment and the degree of curtailment that is made necessary. The information of this table no doubt helps to explain why the birth rate has declined.

Several references have been made to correlating the size of the family, and no doubt the question has arisen in the mind of the reader as to what measure of the size of the family was used. In order to classify according to size of family, it was necessary to find some measure of the size of the family other than the number of persons in the family. A somewhat better unit is that used in analysis of food consumption, in which children and women are considered as various fractions of an adult male; the particular fractions are determined on the basis of the respective calory requirements of the particular individuals. Thus, the Bureau of Labor Statistics rates the adult male as 1.0, the wife as 0.9, child over 14 years as an adult, child from 11 to 14 years as 0.9, child from 7 to 10 years as 0.75, child from 4 to 6 years as 0.4 and child under 3 years as 0.15. This scale is certainly a better unit of measure of the size of the family than the number of persons in the family. Such a scale has been accepted as an index of the size of the

family when studying food requirements and food expenditures. But is it also a good index of the size of family when studying clothing expenditures, rent, expenditures for fuel, light and sundries?

It is possible to find out what is the proportion of expenditure for clothing for the wife and for children of various ages, as compared with the proportion for the adult male. If this is done we can then compare the fractions which the children's and wife's clothing is of the adult male's with the fractions which the children's and wife's calory requirements is of the adult male's requirements. This will furnish an idea of whether the unit of measure of size of family, based on food requirements, is a good measure when we study clothing expenditure. This has been done and has been generalized in such a way as to show on the average what size a family will be when measured in units of clothing expenditure in terms of the adult male as compared with the same family expressed in units of calory requirements in terms of the adult male. These data show a very close correspondence between these two units of measure for size of family, as shown in the equation $X_{19} = 0.81 + 0.85X_3$. Thus when the equivalent adult males in calory requirements are 3, the total clothing expenditure in multiple of the expenditure for the husband's clothes is on the average 3.3, when the equivalent adult males in calory requirements are 4, the corresponding figure in clothing requirements is 4.2, when the equivalent adult males in calory requirements are 5, the corresponding figure in clothing requirements is 5.1, and where the equivalent adult males in calory requirement are 6, the corresponding figure in clothing requirements is 5.9.

This relationship can be expressed in the following terms: when the size of the family, measured in these units based on food requirements, is increased by 1 unit, the average size of the family in units based on clothing expenditures is increased on the average by 0.85 units. We think then that the units of equivalent males, based on food needs, is a good measure of the size of the family, when considering clothing expenditures. These two expenditures include more than 50 per cent. of the total expenditures.

In a recent article in the U. S. Public Health Reports* the authors discuss the unit of the measure of the size of the family and state that a study was made in 1917 among families in cotton mill villages, but that "while the tabulations of these data have not been completed it is indicated that the Atwater scale is roughly indicative of the variations, according to sex and age in the consumption of all articles for which there are individual expenditures." In this article there is also a table computed from data collected by the U. S. Bureau of Labor Statistics in 1911, which shows that the expenditures of other members of the family bear about the same ratios to the husband's expenditure as their calory requirements do to the calory requirements of the husband, thus strengthening the value of the calory scale as a measure of the size of the family. So in this study of the District of Columbia the calory scale of the U. S. Bureau of Labor Statistics was used as a measure of the size of the family.

Just as the correlation coefficients did not yield the information they appeared at first glance to yield, because of the disturbance of other factors, income and size, and it was found desirable to determine the partial coefficients of correlation which showed a relationship when these disturbing factors were kept constant; so the regression equations may be slightly deceptive at first glance, and we need to determine the partial regression equations. For instance from equation, $X_{11} = 3.94 - 0.00053 X_1$, in Table III, we see that an increase of income of \$100 increases the average percentage of expenditure for children's clothing by 0.053 per cent., but it may be that with the size of the family kept constant, the increase in percentage of expenditure for children's clothing for every \$100 increase in income will not be so much. And this is true, for from equation, $X_{11} = -1.17 + 0.000286 X_1 + 1.58 X_3$, in Table VI, an equation in such a form that we can keep the factor, size of family, constant, we see that for every increase in income of \$100 the percentage of expenditure for children's clothing increases only 0.0286 per cent., or about one half what appeared to be the increase when we did not keep the size of the family

* "Disabling Sickness among the Population of Seven Cotton Mill Villages of South Carolina in Relation to Family Income," by Sydenstricker, Wheeler, and Goldberger. Vol. XXXIII, No. 47, November 22, 1918.

constant. These partial regression equations enable us to keep the size of the family constant and tell what percentage of increase (or decrease) accompanies a change of income of any amount. Similarly, they enable us to keep the size of income constant and tell what percentage of increase (or decrease) accompanies a change of any size in the family. These equations then give us the maximum amount of accurate specific information about items of expenditure for various sizes of income and various sizes of family. In other words, we can tell from these equations what a family of any size with an income of any size (between \$800 and \$1,800) spends, on the average, on the various items of expenditure studied in the year 1916 in the District of Columbia.

TABLE VI.
PARTIAL REGRESSION EQUATIONS.

(1)	$X_2 = -82.70 + 0.14921X_1 - 26.31X_3$	(Deficit and surplus)
(2)	$X_4 = 47.00 - 0.01158X_1 + 1.88X_3$	(Food)
(3)	$X_5 = 22.03 - 0.00125X_1 - 6.28X_3$	(Rent)
(4)	$X_6 = 7.00 - 0.00134X_1 + 0.066X_3$	(Fuel and light)
(5)	$X_7 = 5.29 + 0.003407X_1 + 0.449X_3$	(Clothing)
(6)	$X_8 = 19.00 + 0.01008X_1 - 1.94X_3$	(Sundries)
(7)	$X_9 = 3.88 + 0.00134X_1 - 0.50X_3$	(Husband's clothing)
(8)	$X_{10} = 2.69 + 0.00185X_1 - 0.66X_3$	(Wife's clothing)
(9)	$X_{11} = -1.17 + 0.000286X_1 + 1.58X_3$	(Children's clothing)
(10)	$X_{12} = 4.20 - 0.000163X_1 + 0.026X_3$	(Life insurance)
(11)	$X_{13} = 0.41 + 0.00054X_1 - 0.034X_3$	(Religion)
(12)	$X_{14} = 2.95 + 0.00101X_1 - 0.353X_3$	(Furniture)
(13)	$X_{15} = 0.13 + 0.00059X_1 - 0.0062X_3$	(Education)
(14)	$X_{16} = 0.58 + 0.00079X_1 - 0.141X_3$	(Amusement)
(15)	$X_{17} = 4.20 - 0.00076X_1 - 0.456X_3$	(Liquor and tobacco)
(16)	$X_{18} = 4.20 - 0.00033X_1 - 0.233X_3$	(Sickness)
(17)	$X_{19} = 1.01 - 0.00019X_1 + 0.857X_3$	(Equiv. ad. males in clothing ex.)
(18)	$X_{20} = 0.47 + 0.000148X_1 - 0.071X_3$	(Food ex. per man per day)

In regard to errors, the coefficients of correlation less than ± 0.2 are not significant. When the amounts of expenditure for food, rent, etc., are computed from regression equations based on correlations with the actual amounts spent, these amounts thus computed are usually within plus or minus two percent of the respective amounts computed from regression equations based on correlations with percentages of expenditure. The errors of the coefficients in the partial regression equations of Table VI are listed in Table VII. They were derived from the formula,

$$\epsilon_{b_{2.1}} = \frac{\sigma_2(1-r_{1.2}^2)^{\frac{1}{2}}(1-r_{2.3}^2)^{\frac{1}{2}}}{\sigma_1 n^{\frac{1}{2}}(1-r_{1.3}^2)^{\frac{1}{2}}}.$$

The means and standard deviations are listed in Table VIII.

TABLE VII.
ERRORS OF THE REGRESSION COEFFICIENTS OF THE PARTIAL
REGRESSION EQUATIONS.

$\beta_{2.1} = 0.0259$	$\beta_{12.1} = 0.0005$	$\beta_{2.3} = 7.96$	$\beta_{12.3} = 0.154$
$\beta_{4.1} = 0.0012$	$\beta_{13.1} = 0.0002$	$\beta_{4.3} = 0.384$	$\beta_{13.3} = 0.064$
$\beta_{6.1} = 0.0009$	$\beta_{14.1} = 0.0006$	$\beta_{6.3} = 0.278$	$\beta_{14.3} = 0.205$
$\beta_{9.1} = 0.000405$	$\beta_{15.1} = 0.00001$	$\beta_{9.3} = 0.123$	$\beta_{15.3} = 0.00398$
$\beta_{7.1} = 0.0006$	$\beta_{16.1} = 0.000019$	$\beta_{7.3} = 0.196$	$\beta_{16.3} = 0.0595$
$\beta_{8.1} = 0.00014$	$\beta_{17.1} = 0.00042$	$\beta_{8.3} = 0.433$	$\beta_{17.3} = 0.133$
$\beta_{9.1} = 0.0003$	$\beta_{18.1} = 0.0007$	$\beta_{9.3} = 0.110$	$\beta_{18.3} = 0.216$
$\beta_{10.1} = 0.00027$	$\beta_{19.1} = 0.00044$	$\beta_{10.3} = 0.0832$	$\beta_{19.3} = 0.136$
$\beta_{11.1} = 0.00038$		$\beta_{11.3} = 0.116$	

TABLE VIII.
MEANS AND STANDARD DEVIATIONS.

$M_1 = 1145.5$	$M_{11} = 4.55$	$\sigma_1 = 350$	$\sigma_{11} = 2.619$
$M_2 = +1.5$	$M_{12} = 4.11$	$\sigma_2 = 142.10$	$\sigma_{12} = 2.515$
$M_3 = 3.41$	$M_{13} = 0.92$	$\sigma_3 = 1.146$	$\sigma_{13} = 1.072$
$M_4 = 40.14$	$M_{14} = 2.91$	$\sigma_4 = 7.70$	$\sigma_{14} = 3.377$
$M_5 = 19.65$	$M_{15} = 0.78$	$\sigma_5 = 4.548$	$\sigma_{15} = 0.58$
$M_6 = 5.69$	$M_{16} = 1.00$	$\sigma_6 = 2.055$	$\sigma_{16} = 1.017$
$M_7 = 10.72$	$M_{17} = 1.78$	$\sigma_7 = 3.465$	$\sigma_{17} = 2.188$
$M_8 = 23.92$	$M_{18} = 3.03$	$\sigma_8 = 8.136$	$\sigma_{18} = 3.525$
$M_9 = 3.70$	$M_{19} = 3.71$	$\sigma_9 = 1.936$	$\sigma_{19} = 2.416$
$M_{10} = 2.56$	$M_{20} = 0.39$	$\sigma_{10} = 1.67$	$\sigma_{20} = 0.124$

In Table VI, equations, numbers 2, 3, 4, 5 and 6 may be said to express the well-known Engel's laws in mathematical form, as found in the District of Columbia in 1916. Engel's first law is: As the income of a family increased a smaller percentage of it was expended for food. Our equation is quite in conformity and is: $X_4 = 47.00 - 0.01158 X_1 + 1.88 X_3$, where X_3 is the size of the family measured in the scale of calory requirements, and X_1 is annual income, and X_4 is the percentage of expenditure for food. His second law is: As the income of a family increased, the percentage of expenditures for clothing remained approximately the same. The equation, $X_7 = 5.29 + 0.003407 X_1 + 0.449 X_3$, shows that the percentage of expenditure for clothing rises as the income increases. The third law is: With all the incomes investigated, the percentage of expenditure for rent, fuel and light remained invariably the same. Equations number 3 and 4 in Table VI show that for this study the percentage of expenditure for rent falls slightly, as does also and at the same rate the percentage of expenditure for fuel and light. The fourth law was: As the income increased in amount a constantly increasing percentage was expended for education, health, recreation, amusements, etc.

This was also true in the present study as is shown by the equation, $X_8 = 19.00 + 0.01008 X_1 - 1.94 X_3$. These equations may be said to express Engel's laws quantitatively.

Another possible value to these equations is that they make possible comparisons in cost of living in the cities for which they are computed. For instance, the U. S. Bureau of Labor Statistics found in a study of Portland, Me., that certain amounts were on the average spent for food, for clothing, for rent, for fuel and light, and for sundries, these amounts being for an average family of 4.2 equivalent adult males and with an average annual income of \$1,335. Similar amounts of expenditure were found for Boston, but there the average family was 3.8 equivalent adult males and the average annual income was \$1,520. Should it be desirable to make comparisons between the cost of living in the two places from these data, comparisons could not be made because the incomes are of different sizes and the families are also of different sizes. With such equations, however, the amounts of various expenditures for each place could be reduced to a base of the same income and the same sized family, and comparisons should be made as to the cost of living.

The equations also make it possible to tell what a family of any size and of any income (between \$800 and \$1,800) will spend on a variety of items of family expenditure. Thus if a family of any income and any size wanted as a guide an average budget showing distribution of expenditure, it could be easily constructed from the partial regression equations. Thus a family of husband, wife, and children of ages 4, 6 and 8 years, with an income of \$1,155 a year, spent in 1916 on the average \$462 for food, \$227 for rent, \$66 for fuel and light, \$124 for clothing, and \$277 for sundries (including all other expenditures). The husband spent \$43 for clothes, the wife \$30 and for the children there was spent \$52; \$47 was spent for insurance, \$11 for religious purposes, \$34 for furniture and furnishings, \$9 for newspapers, magazines and books, \$12 for such amusements as movies, theatres, pool, and dances, \$20 for liquor and tobacco and \$35 for sickness.

The effect of a growing family without an increase of income can be told quite specifically from these equations. Suppose a husband, wife, and three children aged 1, 3 and 5 years have

not had an increase of income in three years, still remaining at \$1,000, then they will readjust their expenditures in the following manner, assuming they do not go in debt, that there has been no more children and that there has been no substantial change in the price level. They will spend \$25 a year more for food, \$4 less for rent and \$25 less for sundries. The husband will spend \$6 less for his clothing, the woman \$8 less for hers and the children \$19 more for theirs. The amount spent for fuel and light, insurance and religious purposes will be about the same. They will spend \$4 less for furnishings, \$2 less for newspapers and magazines, \$2 less for movies, theatres and similar amusements, \$5 less for liquor and tobacco and \$3 less for sickness.

A few other conclusions of interest may be drawn from the equations in Table VI.

As the income goes down the percentage of expenditure for women's clothing diminishes most, and for children's clothing least, and as the income goes up the percentage of expenditure for women's clothing increases most and for children's least. In this regard women seem to sacrifice most and the children least.

In families of husband, wife and children of ages 4, 6 and 8 years, with an income of \$1,155, the wife's clothing costs 70 per cent. as much as the husband's clothing.

For purposes of comparison with other budgetary studies, it is desirable to state the percentages of expenditure found in this study of the District of Columbia for the year 1916. For a family of husband, wife and children of ages 4, 6 and 8 years, with an income of \$1,155, the following percentages of expenditure were allowed:

Food.....	40 per cent.
Rent.....	19 per cent.
Fuel and light.....	6 per cent.
Clothing.....	11 per cent.
Sundries.....	24 per cent.
Total.....	100 per cent.

An average family of man, wife and three children, of ages 4, 6 and 8 years, in the District of Columbia in 1916, was in debt, if the annual income was less than \$1,155.